

Técnicas de Avaliação

Avaliação: porque?

- Avaliar adequação de *designs*
- Testar a usabilidade e a funcionalidade do sistema
 - Requisitos dos usuários estão efetivamente sendo atendidos?
 - Sistema comporta-se como esperado?
 - Usuário fica satisfeito e consegue realizar suas tarefas?
- Pode ocorrer
 - Em laboratório
 - *In loco*, i.e., no local de uso

Objetivos da Avaliação

- Analisar a cobertura da funcionalidade do sistema
 - Apropriada? Suporta as tarefas necessárias?
 - Atende às expectativas? Permite bom desempenho?
- Analisar o efeito da interface sobre o usuário
 - Usabilidade? Agrada? Fácil de usar?
- Identificar possíveis problemas específicos do sistema
 - Resultados inesperados? Confusão?

Avaliação - Grupos de Métodos

- Inspeção de usabilidade (*predictive evaluation*)
 - Sem envolver usuários e aplicável a qualquer fase do desenvolvimento de um sistema (implementado ou não)
- Testes de Usabilidade
 - Métodos de avaliação centrados no usuário
 - Métodos experimentais ou empíricos
 - Métodos observacionais
 - Métodos consultivos
- Avaliação não deve ser uma fase única no final do processo de *design*
- Método de avaliação deve ser cuidadosamente escolhido

Avaliação - Outros Grupos

- Experimentos Controlados
 - Experimentos controlados cuidadosamente planejados e executados em laboratório
 - Tratamento estatístico para **preparar experimento**
 - Hipótese a ser testada
 - Variáveis de interesse controladas
 - Laboratórios projetados, rigor na observação e monitoramento do uso do sistema
 - Tratamento estatístico para validar os resultados
- Métodos Interpretativos
 - Como o usuário utiliza o sistema em seu ambiente de trabalho, e como esse uso se integra com outras atividades
 - Avaliação participativa, avaliação etnográfica
 - Registro em vídeo/áudio para apoiar a análise

Avaliação: Estudos em Laboratório



- Com os *designers*, sem usuários
- Com usuários (Testes de Usabilidade)
 - Facilidades de registro audio-visual
 - Espelhos de duas faces
 - Computadores instrumentados

Avaliação: Estudos em Laboratório



- Situação artificial: sem interrupção, sem contexto, sem conversas...
 - Interessante quando:
 - Situação de uso real é perigosa ou remota
 - Deseja-se avaliar tarefas restritas
 - Deseja-se deliberadamente manipular o contexto
 - Para comparar designs alternativos em um contexto controlado
 - Detectar problemas/ estudar aspectos específicos
 - Observar procedimentos pouco usados

Avaliação: in Loco

- Designer/avaliador vai ao ambiente de trabalho do usuário para observá-lo em ação com o sistema
- Situação de uso real, 'quase natural'... porém
 - Interrupções, ruído, influência da presença do 'observador', humano ou não...
- Certas tarefas são difíceis de observar em laboratório ou em loco
 - Ex. as muito longas, que consomem dias ou meses...
- Custo é um aspecto relevante
 - Tirar usuário do trabalho vs. montar uma infra-estrutura de observação no local (e talvez atrapalhar...)

Métodos Interpretativos

- Experimental: Formal e objetiva
- Interpretativa: Mais subjetiva
 - Preocupada com usuários
 - Abordagem sociológica e antropológica
- Envolve diretamente os usuários
 - Questionar/observar usuários (formal ou informalmente)

Métodos Interpretativos

- Crítica as limitações de testar hipóteses científicas em um ambiente fechado
 - Lab não é o mundo real
 - Não pode controlar todas as variáveis
 - Contexto é negligenciado
 - Tarefas testadas são artificiais e curtas

Avaliação do *Design*

- Pode ocorrer mais cedo no processo
 - Quanto antes detectar problemas, melhor...
- Muitas técnicas que não envolvem necessariamente usuários
 - Dependem do designer, de um especialista em fatores humanos
- Objetivo: identificar se princípios cognitivos ou resultados empíricos aceitos estão sendo violados
 - Em geral, métodos analíticos...
- Métodos de Inspeção
 - Também podem ser usados para avaliar implementação

Avaliação do *Design*: Inspeção de Usabilidade



- Avaliadores inspecionam ou examinam aspectos relacionados à usabilidade da interface
- Avaliadores
 - Especialistas em usabilidade, consultores de desenvolvimento de software, usuários finais, especialistas em um determinado padrão de interface, ...
- Resultados dependem dos avaliadores
 - Sustentados pela confiança depositada em seus julgamentos
- Métodos variam
 - Quanto a forma de efetuar julgamentos
 - Quanto aos critérios a serem adotados pelos julgadores

Avaliação do *Design*: Inspeção de Usabilidade



- Métodos informais de avaliação (empíricos)
 - Usam a habilidade e a experiência dos avaliadores
- Estudos demonstram que são um bom complemento aos testes de usabilidade com usuários
- Alguns métodos (não mutuamente exclusivos)
 - Percurso Cognitivo (*Cognitive walkthrough*)
 - Avaliação heurística
 - Revisão de *Guidelines*
 - Inspeção de Consistência



Inspeção de Usabilidade: Objetivos

- Encontrar problemas de usabilidade no *design* de uma interface
- Fazer recomendações no sentido de melhorar a usabilidade do *design*
- Usabilidade
 - Facilidade de aprendizado, eficiência de uso, quão 'agradável' ao usuário, frequência de ocorrência e severidade de erros do usuário
- Trabalho de inspeção: identificar, classificar e contar o número de problemas de usabilidade

Inspeção de Usabilidade: Dificuldades



- O que é um 'problema de usabilidade'?
- Muitas vezes um único problema se manifesta de várias formas
 - Um único elemento da interface pode dificultar o aprendizado, torná-lo lento, causar erros, ou simplesmente ser feio/desagradável
- Difícil estabelecer definições, mas...
 - Na maioria dos casos bom senso é suficiente
- Problema de usabilidade
 - Qualquer aspecto do design onde uma mudança pode mudar uma ou mais medidas de usabilidade

Inspeção de Usabilidade



- Identificar os problemas é apenas parte do processo...
- A equipe deve fazer um redesign da interface
 - Tentar corrigir a maior quantidade possível de problemas
 - Tipicamente, os relatórios gerados a partir da lista de problemas contém sugestões...
 - Muitos métodos sugerem reuniões entre a equipe de avaliadores e a equipe de desenvolvimento
 - Discussão de soluções de redesign
 - Custo associado ao redesign
- Problemas devem ser priorizados quanto à sua gravidade
 - graus de severidade: geralmente derivados do impacto causado pelo problema no usuário e no mercado
 - Compromisso: custo vs. severidade

Inspeção de Usabilidade: Métodos

- Avaliação heurística
 - Inspeção da interface tendo como base uma pequena lista de heurísticas de usabilidade
- Percurso Cognitivo (*Cognitive walkthrough*)
 - Avaliador simula um usuário típico 'percorrendo' a interface para executar tarefas típicas
 - + Tarefas frequentes, críticas,

Inspeção de Usabilidade: Métodos

- Revisão de *Guidelines*
 - Interface é analisada para verificar se está de acordo com uma lista de guidelines de usabilidade
 - ~1.000 guidelines, exige *expertise* do revisor
- Inspeção de Consistência
 - Avaliador verifica consistência dentro de uma família de interfaces
 - Terminologia, cores, layout, formatos de entrada e saída, material on line de treinamento e ajuda, ...

Inspeção de Usabilidade: Vantagens



- Métodos facilmente integráveis a processos convencionais de desenvolvimento de software
- Não exigem muita experiência e longo treinamento para que possam ser utilizados
- Boa experiência educacional para designers novatos

Avaliação Heurística

- Adequada tanto para detectar problemas graves como problemas menores
- Adequada para avaliar o design (em estágios iniciais ou avançados)
- Pode ser usada mesmo por quem não tem muita experiência com avaliação
- Também pode ser usada para avaliar implementações

Procedimento

- 1. Obter entradas
- 2. Realizar avaliação independente
- 3. Discussão/Coleta
- 4. Atribuição de taxa de severidade

Percurso Cognitivo

- Revisores avaliam a interface proposta no contexto da execução de uma ou mais tarefas do usuário
- Origem:
 - *walkthrough* para inspeção de código (Engenharia de Software)
 - Percorrer uma sequência de código, passo a passo, detalhadamente, para checar certas características
- Percurso Cognitivo
 - Passos a serem seguidos pelo usuário para executar uma certa tarefa
 - Avaliadores percorrem a sequência de tarefas e ações, passo a passo, para detectar potenciais problemas de usabilidade

Percurso Cognitivo

- Método de Inspeção de Usabilidade
- Foco:
 - avaliar um *design* quanto à sua facilidade de aprendizagem, particularmente aprendizagem por exploração
 - Avaliadores verificam se cada passo é ou não adequado a um usuário novato
 - O usuário seria bem sucedido ao tentar executá-lo?
 - Processo de percurso dividido em duas fases básicas
 - Fase preparatória
 - Fase de análise

Para realizar um percurso cognitivo



É necessário (fase preparatória):

- Uma descrição do protótipo do sistema
 - Pode ser incompleta, mas razoavelmente detalhada
 - Detalhes como posicionamento e termos usados no menu podem fazer enorme diferença
- Uma descrição da tarefa do usuário
 - Deve ser uma tarefa representativa
- Uma lista completa das ações
 - Necessárias para completar a tarefa com o protótipo dado
- Uma descrição dos usuários
 - Quem são e que tipo de experiência e conhecimento os avaliadores podem assumir que eles têm

Para realizar um percurso cognitivo

Fase de análise

- Contar uma 'estória verossímil' sobre como o usuário iria interagir...
- Para cada ação, em cada uma das tarefas, os analistas respondem quatro questões: Os usuários...
 - Farão a ação correta para atingir o resultado desejado?
 - Perceberão que a ação correta está disponível?
 - Irão associar a ação correta ao efeito desejado?
 - Se a ação correta for executada, perceberão um progresso em relação à tarefa?
- Estória verossímil de sucesso
- Estória verossímil de fracasso
 - Se a resposta a alguma das questões acima é negativa

Percurso Cognitivo

- Pode ser efetuado sobre uma especificação da interface
 - Em papel
 - Protótipo não funcional
 - Protótipo funcional
- Pode ser individual ou em grupo
- Grupo pode envolver
 - Outros designers, engenheiros de software, representantes de outras unidades organizacionais (publicidade, treinamento, documentação)

Percurso Cognitivo

- Definindo as entradas
 - Quem são os usuários do sistema?
 - Qual sua experiência e conhecimento técnico?
 - Ex. Usuários de Linux, pessoas que trabalham com o MS Word...
 - Quais tarefas serão analisadas?
 - Todas as que o sistema suporta... ou as mais relevantes... ou as mais problemáticas...
 - Coleção de tarefas deve ser representativa
 - Qual a seqüência de ações correta para cada tarefa?
 - Descrever...
 - Granularidade da descrição depende da expertise do usuário alvo
 - Qual a interface?
 - Como cada tarefa/ação é 'prompted' pelo sistema?
 - Protótipo em papel ou implementação...

Percurso Cognitivo

- Percorrendo as tarefas/ações, respondendo às 4 perguntas...
 - Usuários farão a ação correta para atingir o resultado desejado?
 - Usuário vai saber como iniciar a tarefa?
 - Ex. Sabe o que precisa ser feito para começar a tarefa?
 - Usuários perceberão que a ação correta está disponível?
 - Ex. Opção para disparar a tarefa é claramente indicada no menu?
 - Usuários irão associar a ação correta ao efeito desejado?
 - Ex. Associar um ícone com o que deseja fazer?
 - Se a ação correta for executada, perceberão um progresso em relação à tarefa?
 - Há feedback do que ocorreu?

Percurso Cognitivo

- Questões servem de guia para construir as estórias
 - Exemplos de estórias:
 - DFAB, 11.4.1; Rocha e Baranauskas, Cap. 4
- Importante registrar a informação gerada durante o percurso
 - Anotações, vídeo
- Resultados do percurso podem ser usados para corrigir problemas

Percurso Cognitivo

- Limitações
 - Enfoque em um único atributo de usabilidade
 - Não deve ser usado como único método de avaliação
- Vantagens
 - Detecta conflitos entre designer e usuário quanto à concepção das tarefas
 - Detecta escolhas ruins/inconsistentes de nomes, rótulos, terminologia
 - Detecta respostas inadequadas às ações

Teste de Usabilidade

- Teste com usuário é fundamental!
- Restrições de tempo e recursos
 - ... mas pode reduzir tempo e custos!
- Laboratórios de Usabilidade:
 - Equipe de especialistas em teste e design de interfaces, equipamento para monitoração
 - **Usability Laboratories: A 1994 Survey**
 - <http://www.useit.com/papers/uselabs.html>
 - Microsoft
 - <http://www.microsoft.com/usability/default.htm>
 - Sun
 - <http://www.sun.com/usability/>
 - Empresa que vende equipamentos para labs. de usabilidade
 - <http://www.usabilitysystems.com/>

Testes de Usabilidade

- Não necessariamente requer um laboratório especializado, entretanto...
- Testes devem ser cuidadosamente planejados e preparados
 - Qual o objetivo do teste?
 - Melhorar um design em desenvolvimento, ou...
 - Avaliar a qualidade global de uma interface em fase final de definição

Testes de Usabilidade



- Plano detalhado de teste
 - O que se deseja obter?
 - Quando e aonde vai acontecer?
 - Qual a duração prevista de cada sessão?
 - Qual o suporte computacional e software necessários?
 - Qual o estado do sistema no início do teste?
 - Quem serão os experimentadores?
 - Quem serão os usuários, quantos, como consegui-los?
 - Que tarefas serão solicitadas aos usuários?
 - Qual critério será utilizado para decidir que os usuários terminaram cada tarefa corretamente?
 - Quanto o experimentador poderá ajudar cada usuário?
 - Quais dados serão coletados, como serão analisados?
 - Qual o critério para determinar que a interface é um sucesso?

Testes de Usabilidade

- Problemas: confiabilidade e validade
- Confiabilidade: grau de certeza de que o mesmo resultado será obtido se o teste for repetido
- Validade: resultados do teste refletem os aspectos de interface que se deseja testar
 - Resultados obtidos têm significado fora do laboratório?
 - Cuidados: diferenças individuais entre usuários, escolha de usuários, escolha de tarefas, diferença entre equipamentos
 - Validade requer planejamento cuidadoso e tratamento estatístico adequado dos dados coletados

Testes de Usabilidade

- Escolha de usuários
 - Representativos de usuários reais do sistema
 - Idade, nível educacional,
 - Experiência prévia com uso de computadores, conhecimento do domínio,...
 - Ideal: usuários reais!!
 - Nem sempre é possível...
- Experimentadores
 - Preparação: conhecimento sobre a aplicação, sobre a interface
 - Pode até ser os próprios projetistas, mas isso requer um certo cuidado...

Testes de Usabilidade

- Tarefas
 - Representativas do uso da interface
 - Dar boa cobertura aos componentes mais significativos
 - Poder ser completadas no tempo razoável para uma sessão de teste (1 a 3 horas)
 - Grau de dificuldade gradativa
 - Planejadas para que possam ser interrompidas a qualquer tempo
 - Descrição de cada tarefa a ser efetuada deve ser fornecida por escrito
 - Realista e inserida em um cenário de uso
- Teste piloto
 - 1 a 3 usuários
 - Para refinar os procedimentos definidos
 - Dados obtidos são descartados para efeito de análise

Etapas de um Teste

- Preparação
 - O local, os equipamentos...
- Introdução
 - Explicar aos usuários, colocá-los a vontade, esclarecer objetivos
- Teste
 - Evitar interferir, evitar ajudar
- Sessão Final
 - Ouvir usuários

Etapas de um Teste

- O que os usuários devem saber
 - Propósito do teste é avaliar o sistema, não o usuário
 - Podem expressar suas opiniões livremente
 - Resultados do teste servirão para melhorar a interface
 - O sistema é confidencial
 - Participação no teste é voluntária, e pode ser interrompida por ele
 - Resultados do teste não são públicos, anonimato dos participantes é garantido
 - Caso estejam sendo feitas gravações em vídeo/áudio, explicar
 - Explicar que pode perguntar, mas nem sempre o experimentador pode responder
 - Instruções específicas sobre o teste

Testes de Usabilidade

- Gravações em vídeo podem ser um recurso valioso para avaliação posterior
 - Ideal é não identificar as pessoas
- Entretanto, a análise é difícil
 - Importante complementar registro com anotações, log files...
- Protocolo *Think-aloud*
 - Atmosfera informal e agradável

Testes de Usabilidade

- Pensando em voz alta (**THINK ALOUD**)
 - Usuário verbaliza o que está pensando enquanto usa o sistema
 - Expectativa é que os pensamentos mostrem como o usuário interpreta cada item da interface
 - Inadequada quando o objetivo é obter medidas de desempenho
 - Usuários tendem a ficar mais lentos e cometer mais erros
 - Requer experimentador bem-preparado
 - Estimular o usuário a falar
 - Não interferir no uso do sistema
 - Vantagem: mostra *o que* o usuário está fazendo e *porque* está fazendo, *enquanto* está fazendo
 - Boa estratégia: usuários trabalhando aos pares
 - Outra alternativa: pedir que os usuários comentem depois suas ações gravadas em vídeo

Testes de Usabilidade

- Medidas quantitativas de desempenho
 - Importantes
 - Para avaliar se objetivos de usabilidade foram atendidos
 - Certos objetivos podem ser avaliados por medidas quantificáveis
 - Para comparar produtos competitivos
 - Para pesquisas em fatores humanos
 - Ex. Eficiência de uso (uma guideline de usabilidade)
 - Medir tempos para executar tarefas...

Medidas Típicas Quantificáveis

- Tempo que o usuário gasta para executar uma tarefa
- Número de tarefas completadas em um intervalo de tempo
- Razão entre interações de sucesso e de erro
- Número de erros do usuário
- Número de ações errôneas imediatamente subsequentes
- Número de comandos distintos utilizados pelo usuário
- Número de comandos nunca utilizados
- Frequência de uso do help ou manuais, tempo de consulta
- Quantas vezes o manual resolveu o problema do usuário
- Proporção entre comentários do usuário favoráveis e críticos
- Quantidade de 'tempo morto'
- Número de vezes que o usuário desviou do objetivo da tarefa
- ...

Testes de Usabilidade

- Em geral, um único teste coleta um pequeno sub-conjunto de medidas
- Testes de campo
 - Sistemas colocados em ambientes de uso reais
 - Coleta de dados automática pelo sistema
- Observação síncrona ou remota